



# Docking-enabled pharmacophore model for histone deacetylase 8 inhibitors and its application in anti-cancer drug discovery

Thangapandian Sundarapandian<sup>1</sup>, John Shalini<sup>1</sup>, Sakkiah Sugunadevi, Lee Keun Woo\*

Division of Applied Life Science (BK21 Program), Environmental Biotechnology National Core Research Center (EB-NCRC), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Gyeongsang National University (GNU), 900 Gazwa-dong, Jinju 660-701, Republic of Korea

## ARTICLE INFO

### Article history:

Received 19 April 2010  
Received in revised form 19 July 2010  
Accepted 20 July 2010  
Available online 3 August 2010

### Keywords:

Pharmacophore  
Histone deacetylase 8  
Cancer  
Virtual screening  
Lipinski's rule  
GOLD

## ABSTRACT

Zinc-dependent histone deacetylase 8 removes the epsilon-acetyl groups present in the N-terminal lysine residues of histone proteins, thereby restricting various transcription factors from being expressed. Inhibition of this enzyme has been reported to be a novel strategy in cancer treatment. To identify novel and diverse leads for use in potent histone deacetylase 8 inhibitor design, a pharmacophore model showing high correlation between experimental and estimated activities was generated using the best conformations of training set compounds from molecular docking experiments. The best pharmacophore model was validated using four different strategies and then used in database screening for novel virtual leads. Hit compounds were selected and subjected to molecular docking using GOLD. The top-scored compound was further optimized for improved binding. The optimization step led to a new set of compounds with both improved binding at the active site and estimated activities. The identified virtual leads could be used for designing potent histone deacetylase 8 inhibitors as anti-cancer therapeutics.

© 2010 Elsevier Inc. All rights reserved.

## 1. Introduction

Acetylation is a post-translational modification that is regulated by two functionally opposite enzymes: histone deacetylase (HDAC), a zinc-dependent metalloenzyme, and histone acetyl transferase (HAT) [1,2]. Whereas HDAC enzymes are found in bacteria, fungi, plants and animals, HATs are only present in eukaryotic systems [3,4]. HDACs mediate the removal of epsilon-acetyl groups of lysine residues present at the N-terminal part of core histones. This deacetylation increases the positive charge density on the N-termini of histones, thereby leading to tight histone–DNA (which has negative charge) binding, which restricts the access of transcription factors. By contrast, HATs loosen the histone–DNA binding by acetylating the positively charged lysine residues in the N-termini of core histones, thus activating gene transcription. The catalytic functions of these two enzymes are well balanced in normal cells, but the disturbance of this balance is frequently observed in human cancers [3,5]. Inhibition of HDAC enzymes is an emerging therapeutic strategy for treating cancers in humans [6–8]. Several *in vitro* studies have shown the influence of HDAC inhibitors over proliferation, apoptosis and promotional effects on differentiation in several cancer cell lines; moreover, these inhibitors have exerted

antitumoral effects in mouse models [9,10]. In addition to cancer, other therapeutic applications on neurodegenerative diseases and inflammation have been proposed for HDAC inhibitors [11]. Blocking the activity of HDAC enzymes should favor chromatin relaxation and increase gene transcription, which could be a common mechanism for the induction of gene expression by HDAC inhibitors [3]. HDAC inhibitors are structurally distinct, comprising hydroxamic acids, cyclic peptides, electrophilic ketones, short-chain fatty acids, and benzamides [12–22]. Three structural characteristics are commonly present in all HDAC inhibitors: a zinc-binding group (ZBG) to coordinate the catalytic metal ion, and a hydrophobic spacer and hydrophobic cap group to bind with residues at the tunnel and active site entrance, respectively (shown in Fig. 1). Among the various classes of HDAC inhibitors, hydroxamic acid inhibitors are the most well-studied group and includes Vorinostat (Suberoanilide hydroxamic acid, SAHA), which was recently approved by FDA for the treatment of cutaneous T-cell lymphoma. A handful of molecules from this class are in various phases of drug development. Panobinostat, Belinostat, ITF-2357 are some of the hydroxamic acid derivatives in phase II development. Several other inhibitors, MS-275 and MGCD0103 (benzamide), depsipeptide (cyclic peptide), and valproic acid and butyrate (short-chain fatty acid), are also in active development [23,24]. Almost 30 years ago, some studies to understand the effect of dimethyl sulfoxide on terminal differentiation of murine erythroleukemia cells [25] led to the development of novel pharmacological agents in the area of chromatin remodeling. HDAC enzymes possess a narrow and

\* Corresponding author. Tel.: +82 55 751 6276; fax: +82 55 752 7062.  
E-mail address: [kwlee@gnu.ac.kr](mailto:kwlee@gnu.ac.kr) (L.K. Woo).

<sup>1</sup> These authors have equally contributed to this work.

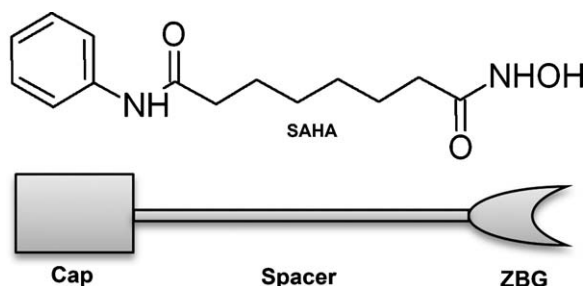


Fig. 1. Basic pharmacophore model of HDAC inhibitors compared with SAHA.

deep active site spanning a length equivalent to four or six straight carbon chains. A divalent cation ( $Zn^{2+}$ ) is located at the bottom of the active site and acts as catalytic machinery in combination with charge relay system residues H142, H143, D176 and D183 (shown in Fig. 2). Recently determined crystal structures of HDAC8 with different set of inhibitors elucidated the binding mode of HDAC8 inhibitors at this active site. The family of HDAC enzymes comprises 18 isoforms, which are broadly classified into four classes. Class I (HDAC1–3 and 8), class II (HDAC4–7, 9 and 10) and class IV (HDAC11) HDACs are zinc-dependent enzymes. Class III HDACs are sirtuin proteins that are dependent on  $NAD^+$  for their activity [26,27]. Most of the currently known inhibitors are not selective to inhibit individual HDAC enzymes because of the highly conserved active site in all HDAC isoforms. Except HDAC8, functional HDACs are not found as single peptides but as multimeric complexes of higher molecular weight; in addition, most of the purified HDAC enzymes are functionally inactive [10,28]. Along with this advantage, expression of HDAC8 notably correlates with the disease stage of neuroblastoma, a highly malignant childhood cancer derived from the sympathetic nervous system [29,30]. Moreover, an RNA interference study showed that HDAC8 is involved in the regulation of proliferation, clonogenic growth and neuronal differentiation of neuroblastoma cells. Inv1, an abnormal fusion protein formed during acute myeloid leukemia binds HDAC8, is also associated with aberrant, constitutive genetic repression [31]. Therefore, HDAC8 is considered to be the best model among other mammalian HDACs from a structural biology perspective. Previously reported pharmacophore models were based on non-specific activity values [32] and HDAC1 inhibitory activities [33], but not for HDAC8.

In our study, we successfully used pharmacophore modeling, virtual screening and molecular docking approaches in the identification of potential virtual leads for HDAC8 inhibitors to treat

various cancers in human. A statistically significant pharmacophore model was developed and validated using four different methods. The validated pharmacophore model was then used in database screening and molecular docking studies to identify the virtual lead candidates for HDAC8 inhibitors. The final hit compounds were optimized to improve the binding at the HDAC8 active site. Finally, the compounds with the optimized substitutions were checked for their synthetic feasibility and novelty.

## 2. Materials and methods

### 2.1. Collection of HDAC8 inhibitors

A sufficiently large set of compounds with their HDAC8 inhibitory activity data is essential in pharmacophore model generation. In literature, a variety of assay procedures are used to determine the HDAC8 inhibition of chemical compounds [34–38]. More than 500 compounds along with their experimental activity data were identified from various scientific resources. Out of these, 97 compounds were found to have  $IC_{50}$  values identified using the same biological assay conditions [38–44]. The  $IC_{50}$  value is the concentration of a compound required to inhibit 50% of the HDAC8 activity.

### 2.2. Training set, test set selection and optimization

The selection of a suitable training set is the most important step in pharmacophore modeling, as this determines the quality of the generated pharmacophores. The test set, which is not used in model generation, but in the pharmacophore validation process, has equal importance. The  $IC_{50}$  values of these 97 compounds spanned across a wide range from 0.008  $\mu M$  to 35  $\mu M$ . In this study, 20 of 97 compounds were chosen as the training set based on the diversity observed in chemical structures and experimental activity values. The remaining 77 compounds were employed in the validation process as the test set. Fig. 3 shows the chemical structures of the training set compounds. Both the training and test sets are sufficiently diverse in their chemical structures. The 2D chemical structures of all the compounds were drawn using ChemSketch, version 12 (Advanced Chemistry Development Inc., Toronto, Canada) and subsequently exported to Accelrys Discovery Studio, version 2.1 (DS) (Accelrys Inc., San Diego, USA). The 3D structure of every compound was generated, followed by optimization using force fields in CHARMM [45]. All of the structures were minimized using *Steepest Descent* followed by *Conjugate Gradient* algorithms

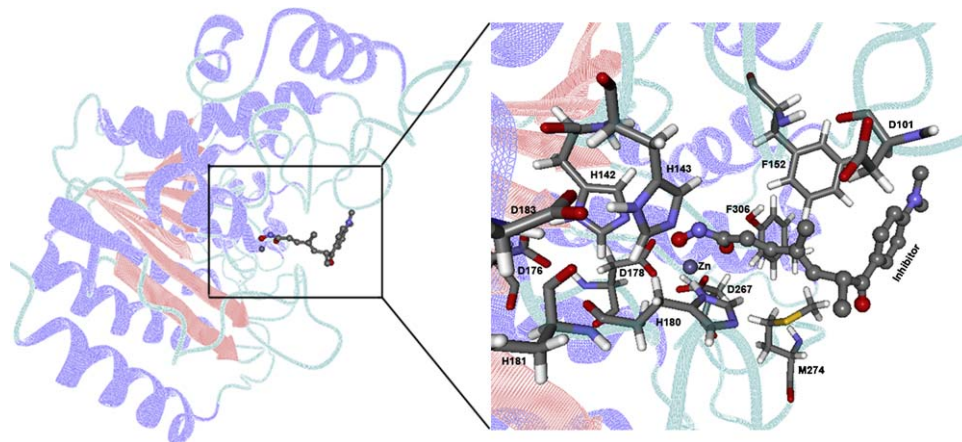


Fig. 2. Full view of the 3D structure of HDAC8 enzyme with one of its hydroxamate inhibitors (left). The active site of HDAC8 enzyme is zoomed in for a clear view (right). The active site residues involved in deacetylation mechanism are labeled. Protein is shown in stick model whereas the inhibitor and metal ion are shown in ball and stick representation.

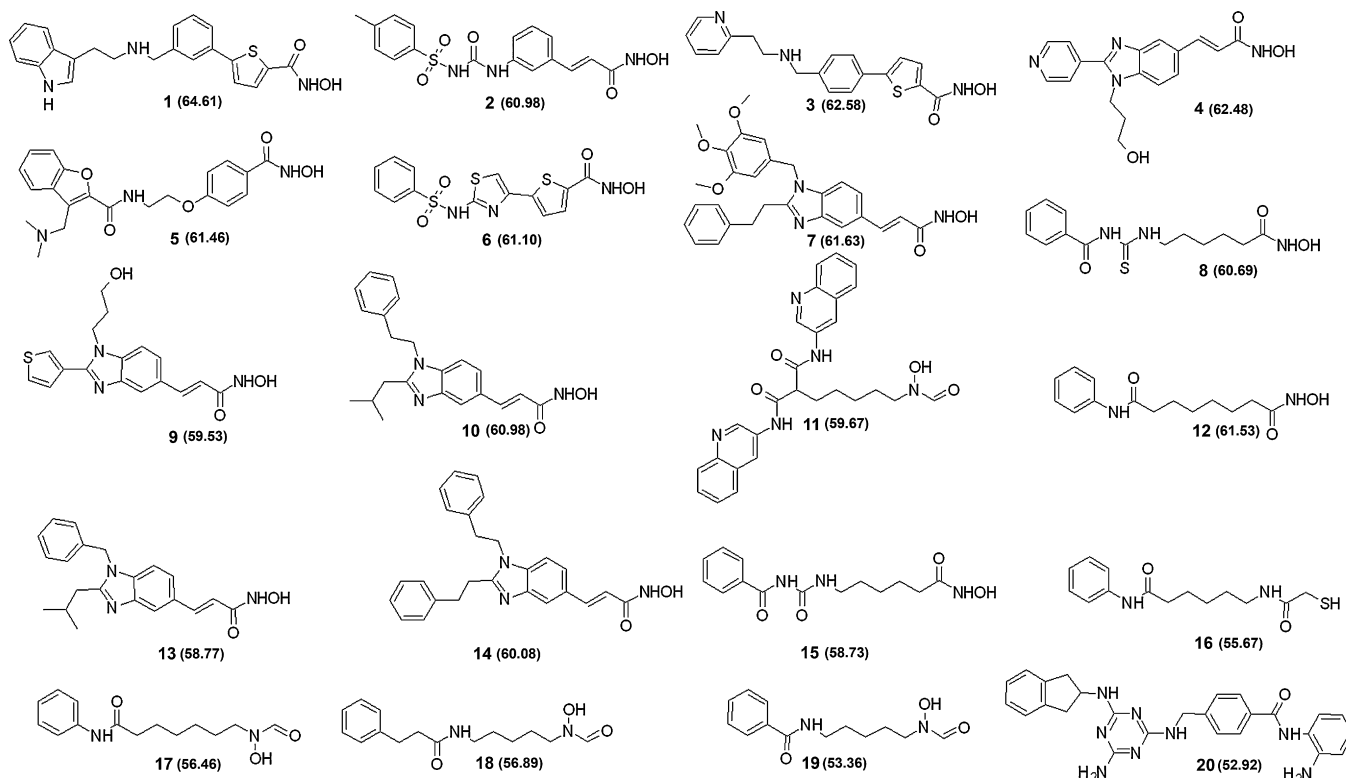


Fig. 3. Training set compounds with their top GOLD fitness scores.

with a convergence gradient value of 0.001 kcal/mol. The activity values of the data set were classified into four categories, active ( $IC_{50} \leq 0.1 \mu M$ , ++++), moderately active ( $0.1 \leq IC_{50} \leq 1 \mu M$ , +++), less active ( $1 \leq IC_{50} \leq 10$ , ++), and inactive ( $IC_{50} > 10$ , +), to simplify the results of pharmacophore generation and validation.

### 2.3. Molecular docking enabled pharmacophore modeling

Every pharmacophore modeling study that employs the *HypoGen* or *HipHop* module of the CATALYST program conventionally starts with the diverse conformation generation step that uses the *ConForm* module with the Poling algorithm of CATALYST to produce a maximum of 255 diverse conformations with a cutoff value of 20 kcal/mol from the local energy minimum [46–48]. Though this step has performed well in various applications, numerous previous studies have revealed that the energy of the biological conformer of a particular compound is usually well above its local energy minimum [49,50]. The well-known induced-fit theory also indicates that a molecule should rearrange itself to fit into the protein's active site; the energy spent for rearrangement is compensated by protein–ligand binding [51,52]. Therefore, no protocol including *ConForm* of CATALYST could assure the prediction of the biological conformation of small molecules. This problem of conformation generation is especially encountered in compounds with higher molecular weights and more rotatable bonds. To avoid this problem, we performed molecular docking experiments and used the best-docked conformations of each training set compound for pharmacophore generation. All of the training set compounds were subjected to molecular docking using GOLD (Genetic Optimization for Ligand Docking), version 4.1 [53]. GoldScore scoring function was chosen out of the other functions such as ChemScore, Astex Statistical Potential (ASP) or Piecewise Linear Potential (PLP) scoring functions. The thirty best binding conformations were chosen based on their GOLD fitness scores and their binding orientations relative to experimentally known conformations of co-crystallized

small molecule inhibitors in various HDAC8 crystal structures. The GOLD program performed well on the training set compounds, as the GOLD fitness scores were well correlated with the activity values (Table S1). The training set compounds are shown with their best GOLD fitness scores in Fig. 3.

Pharmacophore model generation was carried out using the top 30 binding conformations of each training set compound. All of the calculations were performed by the DS program. *Feature Mapping* in DS was used to identify the chemical features present in the training set compounds. Pharmacophore model generation was carried out by selecting chemical features, such as hydrogen bond acceptor (HBA), hydrogen bond donor (HBD), hydrophobic (HYP), positive ionizable (PI) and ring aromatic (RA). *HypoGen* generates pharmacophore models primarily based on the chemical features of the most active compounds in the training set. The default *Minimum Inter-feature Distance* value of 2.97 Å was changed to 2.2 Å so the program would consider the chemical features present at least within 2.2 Å during pharmacophore generation. This change was essential because the hydroxamic acid moiety of HDAC inhibitors, which is necessary to interact with the metal ion in the active site, has a secondary amino and a hydroxyl group within 2.2 Å. The *Uncertainty* value was changed to 2 from the default value of 3 as the training set compounds barely spanned the required range of activity (i.e., four orders of magnitude) [32]. The *Uncertainty* value of 2 is defined by CATALYST as a measured value being two times higher or two times lower than the true value. All other parameters were kept at their default values. The 3D QSAR *Pharmacophore Generation* protocol (*HypoGen* protocol of CATALYST) in DS was used to generate ten pharmacophore models. During hypothesis generation, the structure and activity correlations in the training set were rigorously examined. *HypoGen* identifies features common to the active compounds and excludes features common to the inactive compounds within conformationally allowable regions of space. It further estimates the activity of each training set compound using regression parameters. The parameters were computed by regres-

sion analysis using the relationship of geometric fit value versus the negative logarithm of activity. The better the geometric fit the greater the activity prediction of the compound. The fit function checks if the feature is mapped, and it also contains a distance term, which measures the distance separating the feature on the molecule from the centroid of the feature in the pharmacophore hypothesis. Both terms are used to calculate the geometric fit value.

#### 2.4. Pharmacophore validation

In general, pharmacophore models are used as 3D queries to search chemical databases to identify new and highly potent drug leads. These pharmacophore models should be statistically significant, able to predict the activities of new chemical compounds and retrieve active compounds from the database. The selected pharmacophore model was validated using four methods, cost analysis, test set prediction, Fischer randomization test and enrichment factor calculation (*E*). *HypoGen* ranks the 10 generated pharmacophore models on their cost values. The weight, error and configuration costs are three components that represent the overall cost of a pharmacophore model. The value of the weight cost increases in a Gaussian form as this function weighs a model's deviation from the ideal value of two. The error cost value represents the root mean square (RMS) difference between the experimental and estimated activities of the training set compounds. The configuration cost denotes the complexity or the entropy of the conformational space being optimized and is constant for a given data set. The configuration cost should be less than 17 for a good and significant pharmacophore model. The total cost of a pharmacophore model is the sum of these cost components; however, the error cost is the main contributor.

*HypoGen* also calculates two additional costs for each pharmacophore generation calculation and a cost for every created pharmacophore model. The fixed cost is the lowest possible cost representing a simplest hypothetical model that fits all data perfectly and the null cost represents the maximum cost of a pharmacophore with no features and estimates the activity to be the average activity of the training set compounds and the total cost for every pharmacophore. A larger difference between the fixed and null costs than that between the fixed and total costs signifies the quality of a pharmacophore model. All of these cost values are reported in bits and a difference of 40–60 bits between the total and null costs suggests a 75–90% chance of representing a true correlation in the data. Seventy-seven diverse compounds were used as the test set to validate the pharmacophore model. Fischer randomization (Cat-Scramble) is another approach for pharmacophore model validation. The 95% confidence level was selected in this validation study and 19 random spreadsheets were constructed. This validation method checks the correlation between the chemical structures and biological activity. This method generates pharmacophore hypotheses using the same parameters as those used to develop the original pharmacophore hypothesis by randomizing the activity data of the training set compounds. The fourth validation method is based on the *E* value, which is calculated using a database containing active and inactive compounds.

#### 2.5. Database searching and Lipinski prediction

The selected quantitative pharmacophore model generated within DS was used as a 3D query in database searching. This virtual screening was conducted to find novel and diverse virtual leads suitable for further development. Database searching offers the advantage that the retrieved compounds are usually more easily available for testing than those based on de novo design methods [54]. A molecule must be able to map all of the features of the pharmacophore model to be listed as a hit. All screening exper-

iments were performed using the *Ligand Pharmacophore Mapping* protocol with the *Best Flexible Search* option as available in DS. Hit compounds from the database searching with less than 0.1  $\mu\text{M}$  estimated activity values were retained. In addition, hit compounds with good estimated activity were predicted for the drug-likeness using Lipinski's rule of five [55]. A Lipinski-positive compound has (i) a molecular weight less than 500; (ii) less than 10 hydrogen bond acceptor groups; (iii) less than 5 hydrogen bond donor groups and (iv) an octanol/water partition co-efficient (*Log P*) value less than 5.

#### 2.6. Molecular docking

Compounds that were predicted to be positive in Lipinski drug-likeness screening were subjected to molecular docking studies. The *GOLD* program from Cambridge Crystallographic Data Centre, UK uses a genetic algorithm to dock the small molecules into the protein active site. *GOLD* allows for a full range of flexibility for the ligands and partial flexibility of the protein. Protein coordinates from the crystal structure complex of HDAC8 with trichostatin (PDB ID: 1T64), one of the most active inhibitors, determined at a resolution of 1.9 Å were used to define the active site. The active site was defined with a 10 Å radius around the bound inhibitor. The 10 top-scoring conformations of every ligand were saved at the end of the calculation. Early termination option was used to skip the genetic optimization calculation when any five conformations of a particular compound were predicted within an RMS deviation value of 1.5 Å. The *GOLD* fitness score is calculated from the contributions of hydrogen bond and van der Waals interactions between the protein and ligand, intramolecular hydrogen bonds and strains of the ligand [53,56]. Protein–ligand interactions were analyzed using *DS* and *Molegro Virtual Docker* [57].

#### 2.7. Optimization of final hit compound

Various chemical substitutions were made on the final hit compound with top *GOLD* fitness score for improved binding at the active site. Synthetic accessibility of these optimized compounds was checked using *SYLVIA*, version 1.0 [58,59] from the *Molecular Networks* group. The estimation of synthetic accessibility using *SYLVIA* provides values between 1 for compounds that are very easy to synthesize and 10 for compounds that are very difficult to synthesize. In this study, we set a value of 3 for easily synthesizable compounds and 6 for more difficult compounds. The method for calculating synthetic accessibility accounts for a variety of criteria, such as the complexity of the molecular structure, complexity of the ring system, number of stereo centers, similarity to commercially available compounds, and potential for using powerful synthesis reactions. These criteria are individually weighted to provide a single value for synthetic accessibility. The novelty of the final hits and the optimized structures were evaluated using *SciFinder Scholar* [60] and *Pubchem* compound search [61].

### 3. Results and discussion

#### 3.1. Pharmacophore modeling

We used pharmacophore modeling and molecular docking methodologies to develop a pharmacophore model for HDAC8 inhibitors. The model was further used in database screening to find novel and diverse virtual leads for HDAC8 inhibitors. Molecular docking was performed with the training set compounds prior to pharmacophore generation to obtain the experimentally known conformations of the training set compounds. The well-known pharmacophore (Fig. 1) of HDAC inhibitors contains metal binding and hydrophobic parts that are significantly important for enzyme



**Table 1**  
Statistical results of the generated pharmacophore models.

Hypo	Total cost	Cost difference <sup>a</sup>	RMSD	Error cost	Correlation	Features
1	81.011	61.146	0.941	66.916	0.948	HBA HBD HBD HYP
2	82.021	60.136	1.004	68.139	0.939	HBA HBD HBD HYP
3	92.083	50.074	1.426	78.398	0.871	HBA HBD HBD HYP
4	95.708	46.449	1.545	81.929	0.847	HBA HBD HYP RA
5	97.121	45.036	1.563	82.491	0.851	HBA HBD HYP PI
6	97.228	44.929	1.588	83.273	0.838	HBA HBD HBD HYP
7	97.620	44.537	1.609	83.955	0.832	HBA HBD HYP RA
8	98.308	43.849	1.629	84.611	0.828	HBD HYP RA RA
9	99.456	42.701	1.665	85.790	0.819	HBA HBD HYP RA
10	99.643	42.514	1.658	85.553	0.823	HBA HBD HYP RA

Null cost = 142.157; fixed cost = 71.690; configuration cost = 13.057.

<sup>a</sup> Cost difference = null cost – total cost.

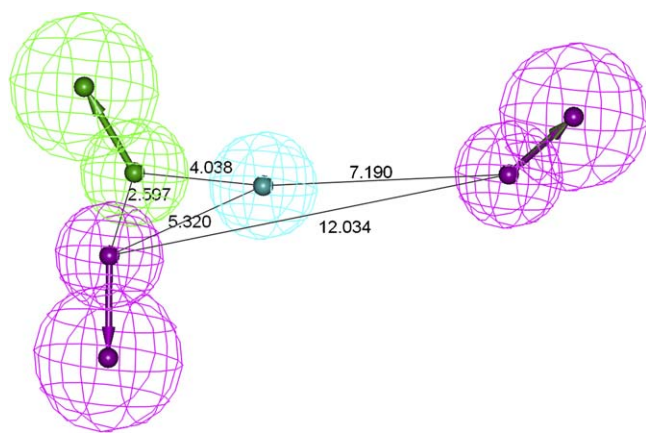
inhibition by binding at the bottom, tunnel and entrance of the active site, respectively. Predicting the correct conformations of HDAC inhibitors is essential for creating the best pharmacophore model. Therefore, pre-molecular docking was performed to find the correct binding orientations, which were then used in quantitative pharmacophore generation. HDAC8, among other HDAC isoforms, is found as a single polypeptide, whereas other isoforms were high molecular weight multimeric complexes. Most of the purified recombinant HDACs are functionally inactive, thus making HDAC8 as the best model among mammalian HDACs [28]. In addition, high-resolution crystal structures of HDAC8 are available with different inhibitors, thereby making this isoform of the HDAC family an interesting target [62,63]. The *HypoGen* protocol in the DS program was employed with the training set compounds for which the inhibitory activity profile is known. Pharmacophore models were generated using HBA, HBD, HYP, PI and RA features as suggested by the *Feature Mapping* protocol, and ten pharmacophore models were exported for further studies. All of the generated pharmacophore models contained at least four chemical features. Four out of ten pharmacophore models had one HBA, two HBD and one HYP features. Other pharmacophore models possessed either HBA or HYP and HBD features along with a PI or RA feature. A significant pharmacophore model should have a large difference

between its total and null cost values. In the present study, the top pharmacophore model (Hypo1) was developed with a cost difference value of 61.146, indicating that the model and data correlated by more than 90%. Further evaluation of the generated pharmacophore models was based on the correlation coefficient. The correlation values of these 10 pharmacophore models were greater than 0.822, and the first two pharmacophore models correlated the activity data with high correlation values, i.e., above 0.9. These results indicate the capability of the pharmacophore model to predict the activity of the training set compounds. Hypo1 showed the highest correlation coefficient value of 0.948, highlighting its strong predictive ability. In addition, RMSD values for the top three pharmacophore models were less than 1.5, further supporting the predictive ability of the top pharmacophore models. Among the ten pharmacophore models, Hypo1 was developed with better statistical values, such as higher correlation, large cost difference, lower RMSD (0.941) and lower configuration cost (13.057). As a result, though it misses the hydrophobic cap group Hypo1 was selected as the best pharmacophore model for further analyses. In order to obtain the compounds with the features identified as in Hypo1 along with the missing hydrophobic cap group, the compounds that are identified via database screening procedure were checked for a hydrophobic group present next to the HBD away from the

**Table 2**  
Comparison of experimental and estimated activity values of the training set compounds based on the best pharmacophore model Hypo1.

Name	Fit value	IC <sub>50</sub> (μM)		Error <sup>a</sup>	Activity scale <sup>b</sup>	
		Experimental	Estimated		Experimental	Estimated
1	7.592	0.008	0.009	1.2	++++	++++
2	7.167	0.022	0.024	1.1	++++	++++
3	6.216	0.120	0.216	1.8	+++	+++
4	6.408	0.161	0.139	-1.1	+++	+++
5	6.243	0.190	0.203	1.1	+++	+++
6	6.250	0.290	0.199	-1.4	+++	+++
7	5.579	0.355	0.936	2.7	+++	+++
8	5.825	0.520	0.531	1	+++	+++
9	5.386	0.537	1.460	2.8	+++	++
10	5.152	0.601	2.503	4.2	+++	+++
11	5.803	0.690	0.558	-1.2	+++	+++
12	5.824	0.820	0.532	-1.5	+++	+++
13	5.471	0.846	1.199	1.4	+++	++
14	5.446	1.005	1.271	1.3	++	++
15	5.530	1.070	1.048	-1	++	++
16	5.355	3.890	1.565	-2.5	++	++
17	5.338	4.000	1.628	-2.4	++	++
18	4.654	9.700	7.871	-1.2	++	++
19	4.682	22.000	7.372	-3	+	++
20	4.369	35.000	15.156	-2.3	+	+

<sup>a</sup> Positive value indicates that the estimated activity is higher than experimental activity and negative value indicates that the estimated activity is lower than experimental activity.<sup>b</sup> HDAC8 enzyme inhibitory activity: active, ++++ (IC<sub>50</sub> ≤ 0.1 μM); moderately active, +++ (IC<sub>50</sub> > 0.1 ≤ 1 μM); less active, ++ (IC<sub>50</sub> > 1 ≤ 10 μM); inactive, + (IC<sub>50</sub> > 10 μM).



**Fig. 4.** The best pharmacophore model Hypo1 represented with distance constraints. Pharmacophoric features colored as follows: hydrogen bond acceptor (green), hydrogen bond donor (magenta), hydrophobic (cyan).

HBA feature. The final hit compounds with no hydrophobic group equivalent to this position are subjected to further optimization by adding variety of hydrophobic groups. The newly identified and designed compounds were investigated for their novelty and synthetic accessibility. Hypo1 consists of one HBA, two HBD and one HYP features. Table 1 shows the statistical parameters of the generated pharmacophore models. Fig. 4 presents the arrangements of the pharmacophoric features along with their inter-feature distance constraints.

### 3.2. Prediction of training set compounds and their overlay on Hypo1

All of the compounds in the training set and test set were categorized into four different groups based on their experimental activity ( $IC_{50}$ ) values: active ( $IC_{50} \leq 0.1 \mu M$ , ++++), moderately active ( $0.1 \leq IC_{50} \leq 1 \mu M$ , +++), less active ( $1 \leq IC_{50} \leq 10$ , ++), and inactive ( $IC_{50} > 10$ , +). The activity of each training set compound was estimated by regression analysis based on the best pharmacophore model Hypo1 and the results are displayed in Table 2. Three out of twenty training set compounds were predicted as having different activities than their experimental values. Noticeably, all active compounds in the training set were predicted as active compounds with estimated activity values of 0.009 and 0.024, which are very close to their experimental  $IC_{50}$  values of 0.008 and 0.022, respectively. Two moderately active compounds were underestimated as less active compounds, and one of the two inactive compounds was overestimated and predicted to be less active. Error values depict the ratio between the experimental and estimated activity values. Positive error values are obtained when the estimated activity value is higher than the experimental value and a negative value indicates the opposite. The first three most active compounds in the training set contained all of the four pharmacophoric features, whereas all of the other compounds mapped three pharmacophoric features of the best pharmacophore model, Hypo1. Fig. 5A and B shows the mapping of the most and least active compounds of the training set on Hypo1, respectively. The carbonyl and the hydroxyl groups present in the hydroxamic acid group of most active compound mapped upon the closely present HBA and HBD features, while the thiophene and secondary amino group located between the phenyl and the indole groups mapped over the HYP and the second HBD features, respectively. The least active compound in the training set only maps with three features of the best pharmacophore model Hypo1. The fit values of these most and least active compounds are 7.592 and 4.369, respectively.

### 3.3. Pharmacophore validation

In addition to the training set prediction, the predictive ability of the best pharmacophore model Hypo1 was evaluated using cost analysis, test set prediction, Fischer randomization, and E value calculation methods. Cost analysis is based on the statistical cost values generated during pharmacophore generation. A diverse test set was used to verify if the pharmacophore model predicts the activity of the compounds that are structurally distinct to the training set. The Fischer randomization test was employed to confirm that the selected model was not generated as a result of chance correlation. The E value calculation was conducted to validate the selectivity of the generated pharmacophore model towards actives rather than inactives.

#### 3.3.1. Cost analysis

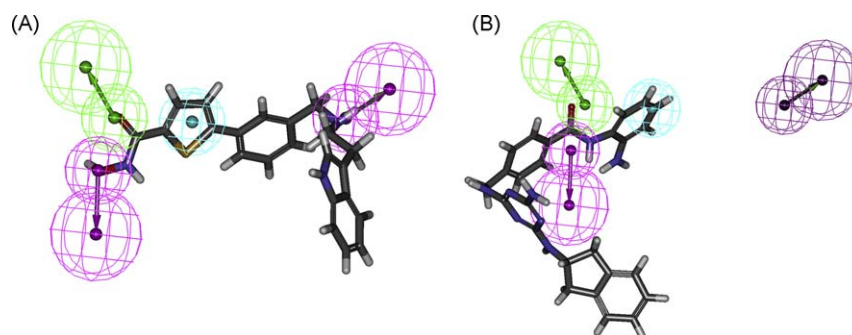
The HypoGen algorithm produced three cost values during pharmacophore generation to assess the quality of the pharmacophore models. The first cost value is the fixed cost value, also known as ideal cost, represents the simplest model that fits the data perfectly. The second one is the null cost value (no correlation cost) represents the highest cost of a pharmacophore with no features estimating the activity to be the average activity data of the training set compounds. A significant pharmacophore should have a large difference between these two cost values. Hypo1 was developed with a fixed cost value of 71.690 and a null cost value of 142.157, thus with a difference of 70.467. The third cost is the total cost value calculated for every pharmacophore model and should be close to the fixed cost value. A large difference between the total and null costs indicates a more significant pharmacophore model. Hypo1 scored a total cost value of 81.011, which is closer to the fixed cost, for a cost difference of 61.146 (shown in Table 1).

#### 3.3.2. Test set prediction

A set of 77 compounds with diverse structures and range of activity values when compared to the compounds in the training set was used to evaluate the best pharmacophore model, Hypo1. The chemical structures of the test set compounds are provided as supplementary material. The Ligand Pharmacophore Mapping protocol with the Best Flexible Search option was used to map all of the test set compounds and the estimated activity values were predicted for each compound. The simple regression between the experimental and estimated activity values of the test set compounds showed a correlation coefficient value of 0.879. In particular, no compound in the test set was predicted with an error value more than 10, thus not exhibiting more than one order of magnitude between experimental and estimated activities (Table 3). Noticeably, 85% (65 compounds) of the test set compounds were predicted within their activity scales while the remaining 15% (12 compounds) were predicted in different activity scales. From these 12 compounds, 3 out of 15 active compounds were underestimated as moderately active; 1 moderately active compound was overestimated as active compound, and 8 moderately active compounds were underestimated as less active compounds. All of the less active and inactive compounds were predicted within their activity scales. Fig. 6 shows the correlation plot for both the training and test set compounds.

#### 3.3.3. Fischer randomization test

The third approach to validate the statistical robustness of the best pharmacophore model is based on Fischer's randomization method. The experimental activities of the training set were scrambled randomly and the resulting training set was used in HypoGen with the parameters chosen for the original pharmacophore generation. A set of 19 random spreadsheets was generated to achieve a 95% confidence level that the best pharmacophore Hypo1 was



**Fig. 5.** Overlay of most active (A) and least active (B) compounds in the training set upon the best pharmacophore model Hypo1. Pharmacophoric features colored as follows: hydrogen bond acceptor (green), hydrogen bond donor (magenta), hydrophobic (cyan). Dark magenta color represents the missing hydrogen bond donor feature. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

not generated by chance. Fig. 7 clearly shows that none of the randomly generated pharmacophore models were produced with better statistical values than Hypo1.

### 3.3.4. Enrichment factor calculation

A database of 1411 compounds (D) including 97 known inhibitors (A) was used in this study. Using the selected pharma-

cophore model, Hypo1, 109 compounds (Ht) were retrieved as hits from the database screening. Among these hits, 92 (Ha) compounds were from the 97 known inhibitors. Therefore, the enrichment factor was calculated to be 12.27, indicating that it is 12 times more probable to pick an active compound from the database than an inactive one. The following formula was used to calculate the E value,  $E = (Ha \times D) / (Ht \times A)$ .

**Table 3**

Test set prediction based on the best pharmacophore model Hypo1.

Name	IC <sub>50</sub> (μM)		Error <sup>c</sup>	Activity scale <sup>d</sup>		Name	IC <sub>50</sub> (μM)		Error <sup>c</sup>	Activity scale <sup>d</sup>	
	Exp <sup>a</sup>	Est <sup>b</sup>		Exp <sup>a</sup>	Est <sup>b</sup>		Exp <sup>a</sup>	Est <sup>b</sup>		Exp <sup>a</sup>	Est <sup>b</sup>
21	0.014	0.015	1.1	++++	++++	60	0.262	0.261	-1	+++	+++
22	0.016	0.016	1	++++	++++	61	0.270	0.271	1	+++	+++
23	0.024	0.181	7.5	++++	++	62	0.290	0.481	1.7	+++	+++
24	0.026	0.035	1.3	++++	++++	63	0.291	1.153	3.9	+++	++
25	0.038	0.059	1.6	++++	++++	64	0.300	0.909	3.1	+++	+++
26	0.040	0.044	1.1	++++	++++	65	0.313	0.313	1	+++	+++
27	0.040	0.033	-1.2	++++	++++	66	0.329	0.92	2.8	+++	+++
28	0.041	0.037	-1.1	++++	++++	67	0.336	0.336	1	+++	+++
29	0.041	0.042	1	++++	++++	68	0.340	3.411	10.1	+++	++
30	0.049	0.051	1.1	++++	++++	69	0.344	0.259	-1.3	+++	+++
31	0.050	0.03	-1.7	++++	++++	70	0.345	0.88	2.6	+++	+++
32	0.051	0.052	1	++++	++++	71	0.350	0.279	-1.3	+++	+++
33	0.072	0.092	1.3	++++	++++	72	0.353	0.975	2.8	+++	+++
34	0.089	0.117	1.3	++++	+++	73	0.366	1.032	2.8	+++	++
35	0.100	0.552	5.5	++++	+++	74	0.400	0.401	1.1	+++	+++
36	0.110	0.093	-1.2	+++	++++	75	0.400	0.585	1.5	+++	+++
37	0.119	0.117	-1	+++	+++	76	0.401	0.392	-1.1	+++	+++
38	0.140	0.141	1	+++	+++	77	0.402	0.4	-1	+++	+++
39	0.140	0.143	1.1	+++	+++	78	0.412	0.944	2.3	+++	+++
40	0.141	0.119	-1.2	+++	+++	79	0.441	0.469	1.1	+++	+++
41	0.145	0.136	-1.1	+++	+++	80	0.454	0.448	-1.3	+++	+++
42	0.148	1.153	7.8	+++	++	81	0.460	0.467	1.2	+++	+++
43	0.150	0.146	-1.1	+++	+++	82	0.650	0.648	-1.1	+++	+++
44	0.159	0.927	5.8	+++	+++	83	0.690	0.688	-1.1	+++	+++
45	0.190	0.252	1.3	+++	+++	84	0.724	0.707	-1.3	+++	+++
46	0.192	1.247	6.5	+++	++	85	0.780	0.770	-1	+++	+++
47	0.200	0.201	1	+++	+++	86	0.790	0.792	1.2	+++	+++
48	0.201	1.311	6.5	+++	++	87	0.790	0.791	1.1	+++	+++
49	0.202	1.032	5.1	+++	++	88	0.800	0.799	-1.1	+++	+++
50	0.203	0.2	-1.1	+++	+++	89	0.883	0.886	1.2	+++	+++
51	0.210	0.217	1	+++	+++	90	1.090	1.083	-1.3	++	++
52	0.230	0.225	-1.1	+++	+++	91	1.450	1.442	-1.2	++	++
53	0.234	0.388	1.7	+++	+++	92	1.710	1.700	-1	++	++
54	0.237	0.586	2.5	+++	+++	93	1.890	1.893	1.1	++	++
55	0.238	0.237	-1	+++	+++	94	2.800	2.640	-1.1	++	++
56	0.250	0.257	1.1	+++	+++	95	6.800	7.289	1.1	++	++
57	0.251	0.25	-1.1	+++	+++	96	7.000	6.852	-1.1	++	++
58	0.254	2.517	9.9	+++	++	97	33.960	36.658	1.2	+	+
59	0.255	0.841	3.3	+++	+++						

<sup>a</sup> Experimental activity.

<sup>b</sup> Estimated activity.

<sup>c</sup> Positive value indicates that the estimated activity is higher than experimental activity and negative value indicates that the estimated activity is lower than experimental activity.

<sup>d</sup> HDAC8 enzyme inhibitory activity: active, ++++ (IC<sub>50</sub> ≤ 0.1 μM); moderately active, +++ (IC<sub>50</sub> > 0.1 ≤ 1 μM); less active, ++ (IC<sub>50</sub> > 1 ≤ 10 μM); inactive, + (IC<sub>50</sub> > 10 μM).

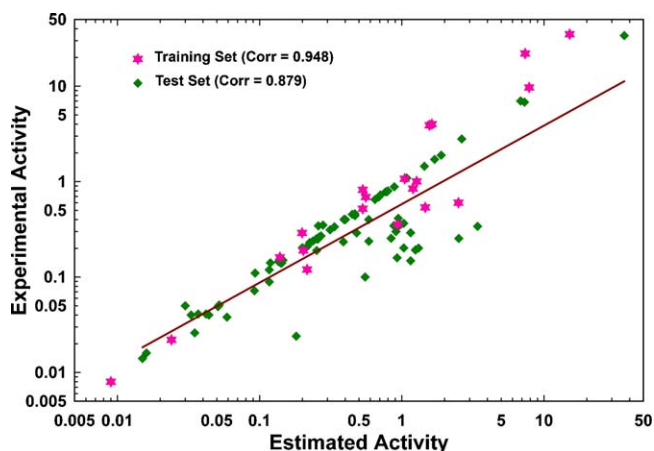


Fig. 6. Correlation plot between experimental and estimated activity values based on Hypo1 for the training and test set compounds.

### 3.4. Database screening and drug-likeness prediction

The best pharmacophore model, Hypo1, was used as a 3D query to search three chemical databases, NCI (260,071 compounds), Maybridge (59,632) and Chembridge (50,000), for a total of 369,703 compounds. The *Ligand Pharmacophore Mapping* protocol with the *Best Flexible Search* option was employed to search these databases. Inhibitory activity values were estimated for the compounds obtained from the database screening. A total of 5342 (4196, 726 and 420 compounds from NCI, Maybridge and Chembridge, respectively) compounds were mapped upon all of the pharmacophoric features present in Hypo1. A total of 1137 compounds scored a *HypoGen* estimated activity value less than 1  $\mu\text{M}$  and were considered for further studies. In addition, violations of Lipinski's rule of five were calculated for the hit compounds and 586 compounds that obeyed Lipinski's rule of five were subjected to molecular docking studies. Fig. 8 lists the steps and outcomes of the database screening procedure.

### 3.5. Molecular docking

All of the final hit compounds were docked into the active site of HDAC8 using GOLD, version 4.1. The active site was defined based on the bound inhibitor in a crystal structure (PDB ID 1T64). The

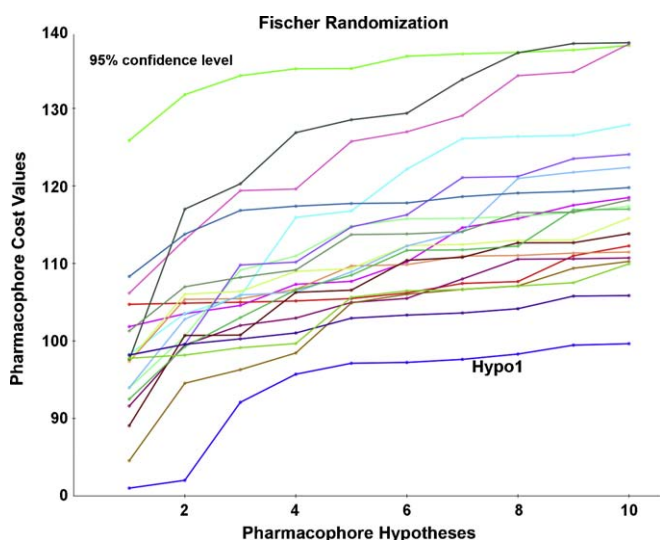


Fig. 7. Results of Fischer randomization test for 95% confidence level.

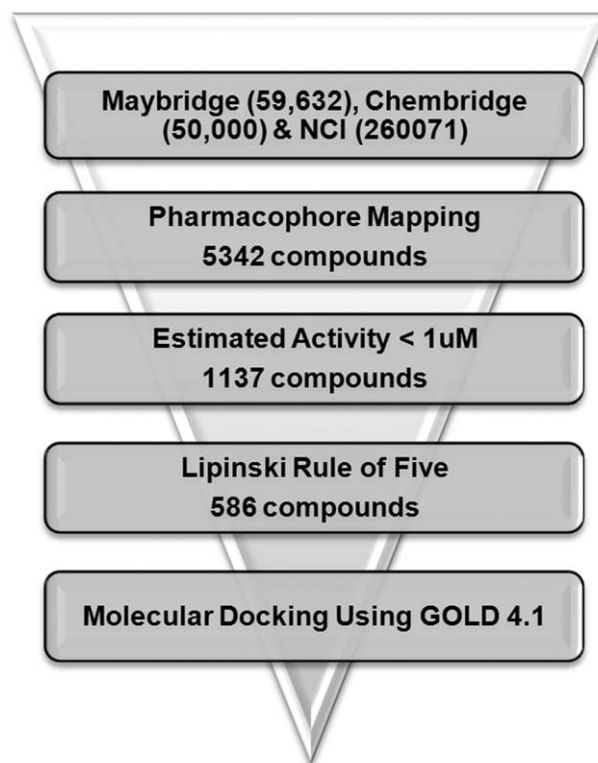
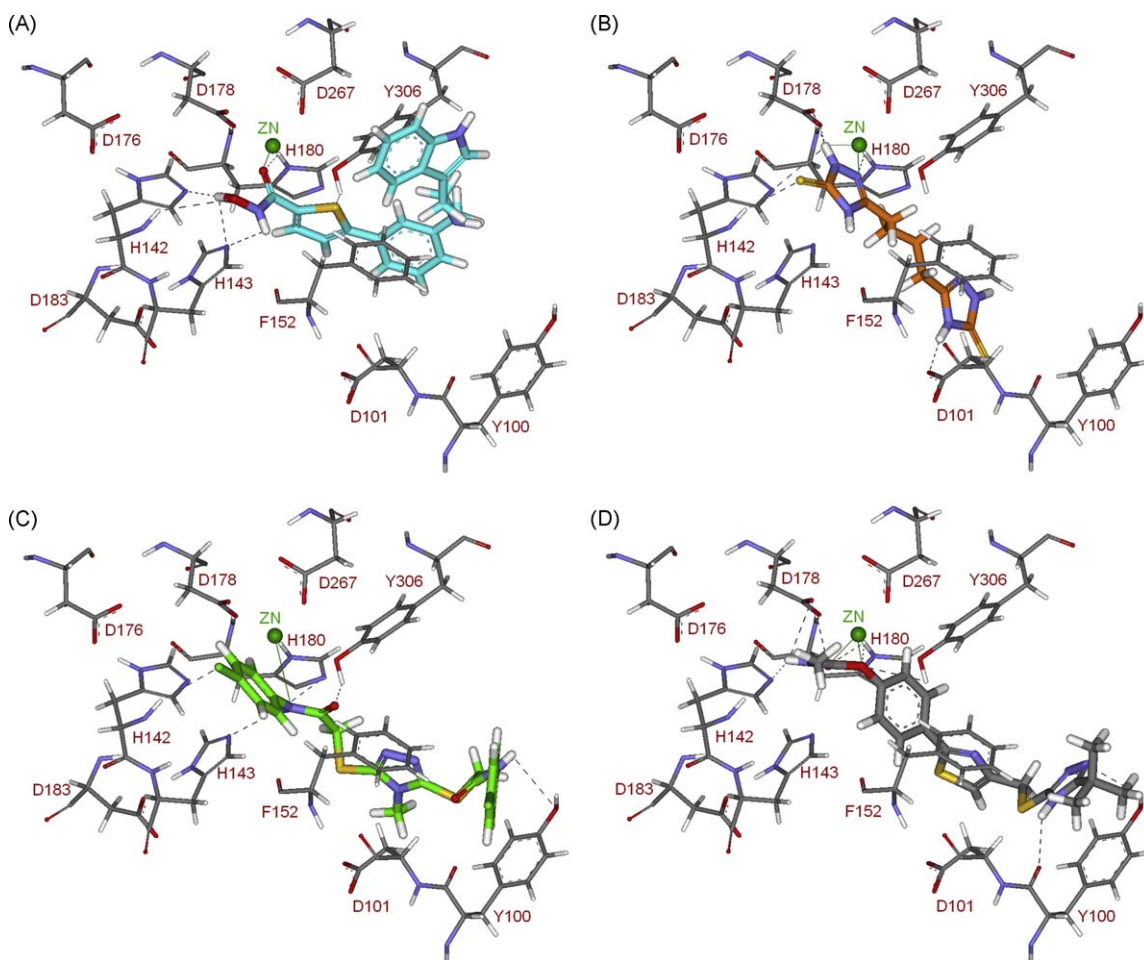


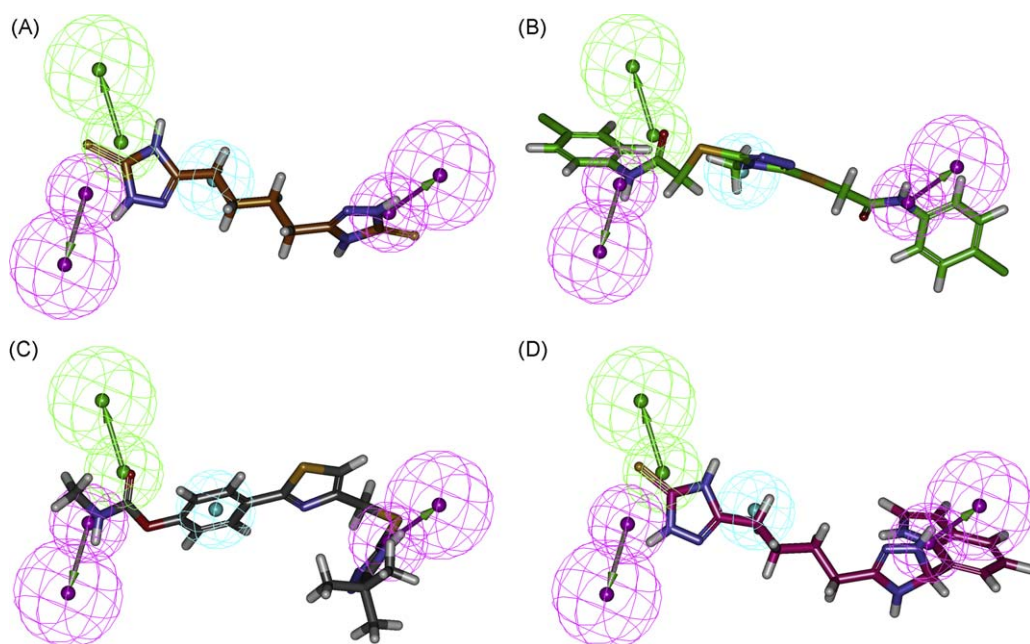
Fig. 8. Results of database screening using Hypo1 as 3D query.

GOLD fitness score along with the binding orientation and hydrogen bond network were considered as the primary components to choose the best poses of the docked compounds. Compound 1 of the training set has scored a GOLD fitness score of 62.566 and was in contact with the catalytically important metal ion and other functionally important active site residues such as H142, H143, H180 and Y306 (shown in Fig. 9A). Binding mode of this compound correlated well with the pharmacophore overlay (Fig. 5A). The hydroxamic acid moiety of this compound mapped on the HBA and HBD features whereas the central thiophene and the only secondary amino groups mapped over the central HYP and second HBD features, respectively. The five-membered thiophene ring fits the hydrophobic tunnel of the active site and the indole ring acts as the hydrophobic cap group for the active site entrance contacts. Prior analyses on the interactions between the various bound inhibitors and HDAC8 in the available crystal structures of HDAC8 revealed that metal ion ( $\text{Zn}^{2+}$ ), H142, H143 and Y306 are the important active site components. Based on this information as well as the GOLD fitness score, the docked poses of the database hit compounds were prioritized. The compounds with GOLD fitness scores above 60 were considered for the next step of the analysis. Sixty-four compounds scored GOLD fitness scores more than 60. Careful visual inspection of the binding orientation and the hydrogen bond network with the catalytic machinery at the active site allowed us to select the compounds we hypothesized would act as novel and potential leads in drug design. Finally, three compounds, namely HTS 09035, BTB 08560 and RF 02863, were identified with the GOLD fitness scores and estimated activity values of 73.224  $\mu\text{M}$ , 68.538  $\mu\text{M}$  and 64.809  $\mu\text{M}$  and 0.807  $\mu\text{M}$ , 0.49  $\mu\text{M}$  and 0.004  $\mu\text{M}$ , respectively. The hydrogen bond network and the metal ion interaction were also observed for these hits. HTS 09035 forms a hydrogen bond network with H142, D178, H180, and D101 and interacts with the metal ion (Fig. 9B). This compound is a derivative of 1,2,4-thiazole-3-thione with a long alkyl chain. The comparison of pharmacophore overlay of this compound with its





**Fig. 9.** Molecular docking results. Docked orientations of (A) compound 1 of training set (cyan color) (B) HTS 09035 (orange color) (C) BTB 08560 (green color) (D) RF02683 (gray color). Active site residues are shown stick form and metal ( $Zn^{2+}$ ) ion is shown in green sphere. Hydrogen bond network with protein residues and metal ion is represented in black dotted and green straight lines, respectively.



**Fig. 10.** The pharmacophore overlay of hit compounds. (A) HTS 09035 (B) BTB 08560 (C) RF 02863 (D) OPT-15.

**Table 4**

List of optimized compounds from HTS 09035 and their estimated activity values with SYLVIA score.

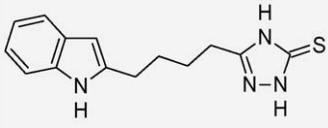
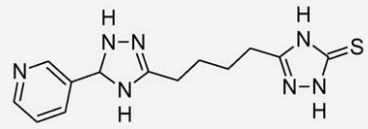
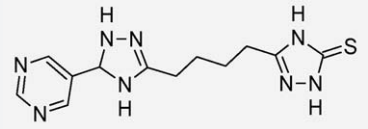
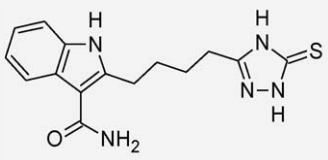

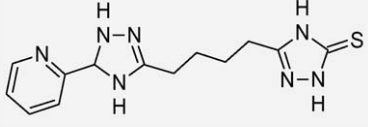
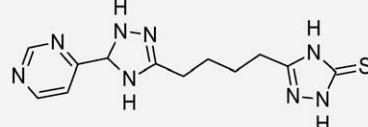
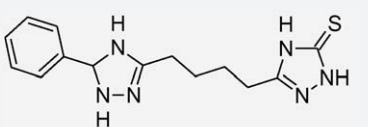
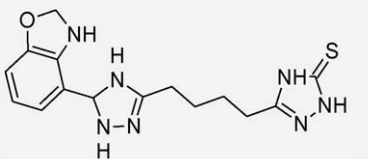
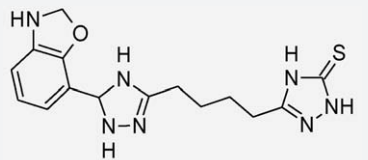
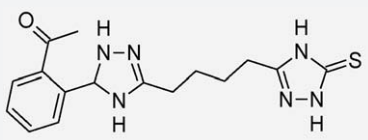
Name	Structure	Estimated Activity	SYLVIA score <sup>a</sup>
OPT1		0.413	4.051
OPT2		0.373	4.675
OPT3		1.201	4.764
OPT4		0.257	4.326
OPT5		0.502	4.14
OPT6		0.583	4.677
OPT7		0.967	4.749
OPT8		0.407	4.445
OPT9		0.442	4.884
OPT10		0.237	4.884
OPT11		0.291	4.949

Table 4 (Continued)

Name	Structure	Estimated Activity	SYLVIA score <sup>a</sup>
OPT12		0.369	4.661
OPT13		0.688	5.139
OPT14		0.381	5.058
OPT15		0.11	5.067
OPT16		0.447	5.211
OPT17		0.153	4.594
OPT18		0.329	4.594
OPT19		0.226	4.658
OPT20		0.184	4.358
OPT21		0.131	4.86

Table 4 (Continued)

Name	Structure	Estimated Activity	SYLVIA score <sup>a</sup>
OPT22		0.165	4.792
OPT23		0.203	4.801
OPT24		0.365	4.957

<sup>a</sup> Synthetic accessibility score where the threshold value 3 indicates the compounds that are easy to be synthesized and value 6 indicates the compounds that are difficult to be synthesized.

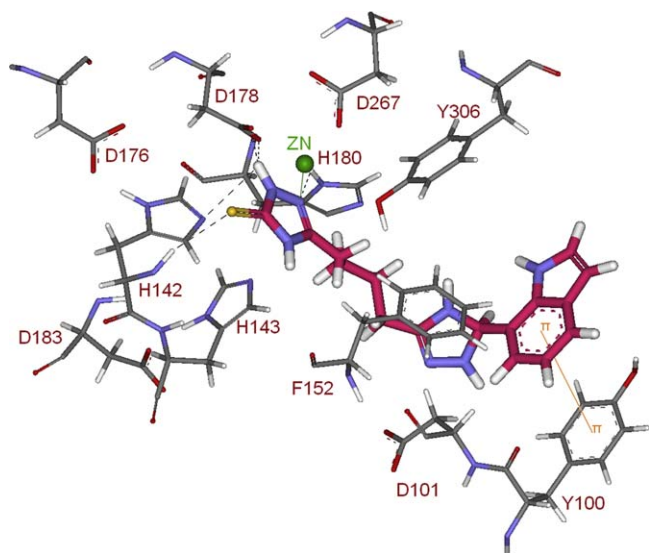


Fig. 11. Docked conformation of the top scored compound after optimization.

molecular docking results revealed that this compound contains all the generated pharmacophoric features of this study but missed the hydrophobic cap group of the general pharmacophore model for HDAC inhibitors. The binding mode of this compound at the active site showed that two five-membered rings that mapped on two

HBD and one HBA features of Hypo1 formed contacts to the catalytic machinery that includes a metal ion, two histidine and aspartate residues and the residues at the entrance. The long alkyl chain that mapped the central HYP feature of Hypo1 fits the hydrophobic tunnel of the active site (Fig. 10A). As this compound missing the hydrophobic cap group it was subjected to further optimization. BTB 08560 shows hydrogen bond interactions with H142, H143, H180, Y306, and Y100 and coordinates the metal ion (Fig. 9C). It is a derivative of N-phenylacetamide derivative with a five-membered ring in the center. The binding mode and pharmacophore overlay of this compound showed that the amide moieties that mapped well on the two HBD and one HBA features of Hypo1 have formed polar contacts in the active site while the central five-membered methyl triazole ring was located at the tunnel enabling hydrophobic contacts (Fig. 10B). RF 02863 also forms hydrogen bond interactions with Y100, D101, H142, D178, H180, and Y306 and also binds to the metal ion (Fig. 9D). This compound is a derivative of phenylcarbamate. Carbamate part of this compound that mapped over the closely present HBA and HBD features of Hypo1 found in contact with the catalytic machinery residues. The terminal triazole and central phenyl rings mapped over the other HBD and HYP features of Hypo1, respectively, are in contact with the residues present in the entrance and tunnel part of the active site (Fig. 10C). All of these identified hits are of diverse scaffolds and thereby providing wide opportunities for future HDAC8 inhibitor design. The novelty and synthetic accessibility of the hits were assessed using *SciFinder scholar* and *SYLVIA 1.0*, respectively. The *SciFinder* results confirmed that these compounds had not been previously tested for

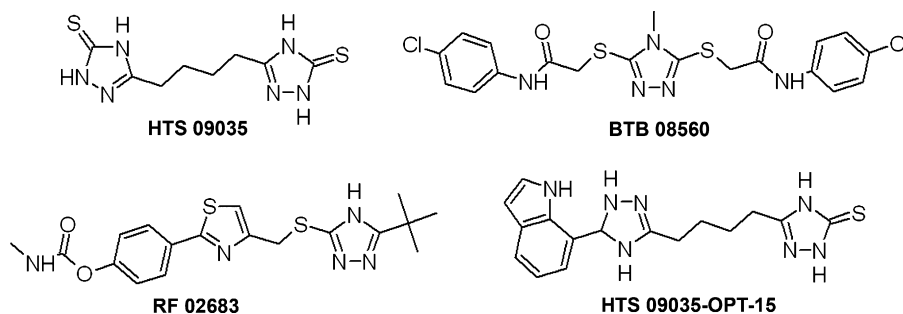


Fig. 12. 2D representations of the final hit compounds.



HDAC8 inhibition. The SYLVIA synthetic accessibility scores for HTS 09035, BTB 08560 and RF 02683 were 3.94, 2.86 and 3.95 respectively. Based on the combination of these results, we suggest that the identified compounds are novel and potent virtual leads for HDAC8 inhibitor design.

### 3.6. Optimization of HTS 09035 for a hydrophobic cap group

All of the final hit compounds bound strongly at the enzyme's active site. Although HTS 09035 scored a high GOLD fitness score, it lacks a hydrophobic cap group, which is one of the essential pharmacophoric points for strong interactions at the active site. This hydrophobic cap group interacts at the entrance of the active site. Various substitutions were made to improve the entrance binding of this hit compound. Table 4 shows the list of modifications made to HTS 09035 and their estimated activity values. A five-membered ring with a thiocarbonyl group at one end that supposed to bind with charge relay system residues at the bottom of the active site was kept constant; the ring at the other end was replaced with various hydrophobic substitutions to improve binding at the active site entrance. Then, we substituted a linear thiourea moiety for the entire five-membered ring at the bottom of the active site to test its significance. Among the various substitutions, a 1-indole substitution at the place of the thiocarbonyl group (OPT-15) exhibited a better estimated activity value of 0.11 than other substitutions or its precursor HTS 09035. The pharmacophore overlay of this compound revealed that it can map all the features of Hypo1 well with the added hydrophobic group, which can bind the residues form the entrance of the active site (Fig. 10D). The thiourea analogue of the same compound scored an estimated activity of 0.203. A compound containing 3,5-pyrimidine at the same position (OPT-3) scored the least estimated activity. Moreover, OPT-15 scored an improved GOLD fitness score of 75.98 and a gained  $\pi$ - $\pi$  interaction with Y100 at the entrance of the active site (Fig. 11). The synthetic accessibility of all of the optimized compounds was checked using SYLVIA, version 1.0 and the results are displayed in Table 4. In addition, 2D representations of the final hits and OPT-15 are shown in Fig. 12. The novelty of the reported hits was also confirmed by SciFinder Scholar and Pubchem compound searches.

## 4. Conclusion

In this work, we built a pharmacophore model with a diverse training set containing 20 compounds with experimental  $IC_{50}$  values for HDAC8 inhibition. The best pharmacophore model was made of one HBA, two HBD and one HYP features with a high correlation value of 0.948 and was validated using 77 structurally distinct test set compounds with a correlation coefficient value of 0.879. In addition, Fischer randomization and *E* value calculation were also used in pharmacophore validation. The validated pharmacophore model was applied as a 3D query in a database search to identify novel scaffolds for use in drug discovery of HDAC8 inhibitors. Five hundred eighty-six compounds with estimated activity values less than 1  $\mu$ M and favorable drug-like properties were selected and docked into the active site of HDAC8. Based on GOLD fitness scores, binding orientations and the interaction network with the catalytic machinery of the deacetylation enzyme, three final lead candidates were suggested for use in HDAC8 inhibitor design. Visual inspection and docking results revealed that the top hit compound, which scored the highest GOLD fitness score, lacks a hydrophobic cap for interacting with the active site entrance. A variety of substitutions were made on the top hit compound to improve the binding at the active site, particularly to enhance entrance recognition. The replacement of the

thiocarbonyl group with an indole improved the binding whereas other substitution either decreased or maintained the binding affinity at the active site. Many of the optimized compounds scored better estimated activity values. The optimized compounds with better estimated activity values were analyzed for their synthetic feasibility and novelty, and the results confirmed their ease of synthesis and novelty. Altogether, the results from all of the experiments provided a set of novel virtual leads for HDAC8 inhibitor design.

## Acknowledgements

This research was supported by Basic Science Research Program (2009-0073267), Pioneer Research Center Program (2009-0081539), and Environmental Biotechnology National Core Research Center program (20090091489) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST). And all the students were recipients of fellowships from the BK21 Program of MEST.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2010.07.007.

## References

- [1] M. Antonello, M. Silvio, R. Rino, E. Monica, S. Gianluca, N. Giuseppina, S. Roberto, J. Florian, L. Peter, B. Gerald, Binding mode analysis of 3-(4-benzoyl-1-methyl-1H-2-pyrrolyl)-N-hydroxy-2-propenamide: A new synthetic histone deacetylase inhibitor inducing histone hyperacetylation, growth inhibition, and terminal cell differentiation, *J. Med. Chem.* 45 (2002) 1778–1784.
- [2] S. Vadivelan, B.N. Sinhab, G. Rambabu, B. Kiran, A.R.P.J. Sarma, Pharmacophore modeling and virtual screening studies to design some potential histone deacetylase inhibitors as new leads, *J. Mol. Graph. Model.* 26 (2008) 935–946.
- [3] P. Marielle, P. Marina, B. Monica, F. Daniela, Histone deacetylase inhibitors: From bench to clinic, *J. Med. Chem.* 51 (2008) 1505–1529.
- [4] J.B. Laura, A.P. Jane, G. Yongli, R.P. Mark, S.H. Charles, T.A. Anthony, Properties of the type B histone acetyltransferase Hat1: H4 tail interaction, site preference, and involvement in DNA repair, *J. Biol. Chem.* 282 (2007) 836–842.
- [5] C. Yadong, L. Huifang, T. Wanquan, Z. Chengchao, J. Yongjun, Z. Jianwei, Y. Qingsen, Y. Qidong, 3D-QSAR studies of HDACs inhibitors using pharmacophore-based alignment, *Eur. J. Med. Chem.* 44 (2009) 1–9.
- [6] P.A. Jones, S.B. Baylin, The fundamental role of epigenetic events in cancer, *Nat. Rev. Genet.* 3 (2002) 415–428.
- [7] C.B. Yoo, P.A. Jones, Epigenetic therapy of cancer: Past, present and future, *Nat. Rev. Drug Discov.* 5 (2006) 37–50.
- [8] S. Minucci, P.G. Pelicci, Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer, *Nat. Rev. Cancer* 6 (2006) 38–51.
- [9] P. Marks, R.A. Rifkind, V.M. Richon, R. Breslow, T. Miller, W.K. Kelly, Histone deacetylases and cancer: Causes and therapies, *Nat. Rev. Cancer* 1 (2001) 194–202.
- [10] J.E. Bolden, M.J. Peart, R.W. Johnstone, Anticancer activities of histone deacetylase inhibitors, *Nat. Rev. Drug Discov.* 5 (2006) 769–784.
- [11] V.K. Galina, F. György, L.G. Jennifer, G.M. Dewey, S. Edward, A.E. Felicia, Phosphorus-based SAHA analogues as histone deacetylase inhibitors, *Org. Lett.* 5 (2003) 3053–3056.
- [12] A.P. Kozikowski, Y. Chen, A. Gaysin, B. Chen, M.A. D'Annibale, C.M. Suto, B.C. Langley, Functional differences in epigenetic modulators—superiority of mercaptoacetamide-based histone deacetylase inhibitors relative to hydroxamates in cortical neuron neuroprotection studies, *J. Med. Chem.* 50 (2007) 3054–3061.
- [13] P.A. Marks, R. Breslow, Dimethyl sulfoxide to vorinostat: Development of this histone deacetylase inhibitor as an anticancer drug, *Nat. Biotechnol.* 25 (2007) 84–90.
- [14] W.S. Xu, R.B. Parmigiani, P.A. Marks, Histone deacetylase inhibitors: Molecular mechanisms of action, *Oncogene* 26 (2007) 5541–5552.
- [15] M. Dokmanovic, C. Clarke, P.A. Marks, Histone deacetylase inhibitors: Overview and perspectives, *Mol. Cancer Res.* 5 (2007) 981–989.
- [16] K.V. Butler, A.P. Kozikowski, Chemical origins of isoform selectivity in histone deacetylase inhibitors, *Curr. Pharm. Des.* 14 (2008) 505–528.
- [17] G. Estiu, E. Greenberg, C.B. Harrison, N.P. Kwiatkowski, R. Mazitschek, J.E. Bradner, O. Wiest, Structural origin of selectivity in class II-selective histone deacetylase inhibitors, *J. Med. Chem.* 51 (2008) 2898–2906.
- [18] N. Khan, M. Jeffers, S. Kumar, C. Hackett, F. Boldog, N. Khrantsov, X. Qian, E. Mills, S.C. Berghs, N. Carey, P.W. Finn, L.S. Collins, A. Tumber, J.W. Ritchie, P.B. Jensen, H.S. Lichenstein, M. Sehested, Determination of the class and isoform

- selectivity of small-molecule histone deacetylase inhibitors, *Biochem. J.* 409 (2008) 581–589.
- [19] O. Moradei, A. Vaisburg, R.E. Martell, Histone deacetylase inhibitors in cancer therapy: New compounds and clinical update of benzamide-type agents, *Curr. Top Med. Chem.* 8 (2008) 841–858.
- [20] S. Shankar, R.K. Srivastava, Histone deacetylase inhibitors: Mechanisms and clinical significance in cancer: HDAC inhibitor-induced apoptosis, *Adv. Exp. Med. Biol.* 615 (2008) 261–298.
- [21] P. Jones, M.J. Bottomley, A. Carfi, O. Cecchetti, F. Ferrigno, P. Lo Surdo, J.M. Ontoria, M. Rowley, R. Scarpelli, C. Schultz-Fademrecht, C. Steinkühler, 2-Trifluoroacetylthiophenes, a novel series of potent and selective class II histone deacetylase inhibitors, *Bioorg. Med. Chem. Lett.* 18 (2008) 3456–3461.
- [22] J. Schemies, W. Sippl, M. Jung, Histone deacetylase inhibitors that target tubulin, *Cancer Lett.* 280 (2009) 222–232.
- [23] P.A. Marks, W.S. Xu, Histone deacetylase inhibitors: Potential in cancer therapy, *J. Cell. Biochem.* 107 (2009) 600–608.
- [24] L. Stimson, V. Wood, O. Khan, S. Fotheringham, N.B. La Thangue, HDAC inhibitor-based therapies and haematological malignancy, *Ann. Oncol.* 20 (2009) 1293–1302.
- [25] P.A. Marks, Discovery and development of SAHA as an anticancer agent, *Oncogene* 26 (2007) 1351–1356.
- [26] C. Monneret, Histone deacetylase inhibitors, *Eur. J. Med. Chem.* 40 (2005) 1–13.
- [27] M. Nobuko, T. Yasunari, K. Masayo, T. Seiji, I. Masamichi, N. Takeo, T. Minoru, Role of NAD-dependent deacetylases SIRT1 and SIRT2 in radiation and cisplatin-induced cell death in vertebrate cells, *Genes Cells* 10 (2005) 321–332.
- [28] A. Vannini, C. Volpari, P. Gallinari, P. Jones, M. Mattu, A. Carfi, R. De Francesco, C. Steinkühler, S. Di Marco, Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex, *EMBO Rep.* 8 (2007) 879–884.
- [29] G.M. Brodeur, Neuroblastoma: Biological insights into a clinical enigma, *Nat. Rev. Cancer* 3 (2003) 203–216.
- [30] I. Oehme, H.E. Deubzer, D. Wegener, D. Pickert, J.P. Linke, B. Hero, A. Kopp-Schneider, F. Westermann, S.M. Ulrich, A. von Deimling, M. Fischer, O. Witt, Histone deacetylase 8 in neuroblastoma tumorigenesis, *Clin. Cancer Res.* 15 (2009) 91–99.
- [31] L.D. Kristie, L. Bart, K. Tanawan, D.F. Alan, W.H. Scott, The inv(16) fusion protein associates with corepressors via a smooth muscle myosin heavy-chain domain, *Mol. Cell. Biol.* 23 (2003) 607–619.
- [32] C. Ya-dong, J. Yong-Jun, Z. Jian-Wei, Y. Qing-Sen, Y. Qi-Dong, Identification of ligand features essential for HDACs inhibitors by pharmacophore modeling, *J. Mol. Graph. Model.* 26 (2008) 1160–1168.
- [33] Y. Liqin, L. Fei, C. Yadong, Y. Qidong, Pharmacophore identification of hydroxamate HDAC 1 inhibitors, *Chin. J. Chem.* 27 (2009) 557–564.
- [34] D.J. Witter, P. Harrington, K.J. Wilson, M. Chenard, J.C. Fleming, B. Haines, A.M. Kral, J.P. Secrist, T.A. Miller, Optimization of biaryl selective HDAC1&2 inhibitors (SHI-1:2), *Bioorg. Med. Chem. Lett.* 18 (2008) 726–731.
- [35] O.M. Moradei, T.C. Mallais, S. Frechette, I. Paquin, P.E. Tessier, S.M. Leit, M. Fournel, C. Bonfils, M.C. Trachy-Bourget, J. Liu, T.P. Yan, A.H. Lu, J. Rahil, J. Wang, S. Lefebvre, Z. Li, A.F. Vaisburg, J.M. Besterman, Novel aminophenyl benzamide-type histone deacetylase inhibitors with enhanced potency and selectivity, *J. Med. Chem.* 50 (2007) 5543–5546.
- [36] L. Tao, K. Galina, A.E. Felicia, Design and synthesis of a potent histone deacetylase inhibitor, *J. Med. Chem.* 50 (2007) 2003–2006.
- [37] C. Yufeng, L. Miriam, N.S. Doris, D.B. Daniel, S.D. Geoffrey, P.K. Alan, A series of potent and selective. Triazolylphenyl-based histone deacetylases inhibitors with activity against pancreatic cancer cells and *Plasmodium falciparum*, *J. Med. Chem.* 51 (2008) 3437–3448.
- [38] G. Wenxin, N. Inna, D.S. Ronald, M.H. Curt, B.S. Richard, Carbonyl-sulfur-containing analogs of suberoylanilide hydroxamic acid: Potent inhibition of histone deacetylases, *Bioorg. Med. Chem.* 14 (2006) 3320–3329.
- [39] Y.H.W. Tom, H. Christian, W. Yiqin, D. Sheng, G.S. Peter, Design, synthesis, and activity of HDAC inhibitors with a N-formyl hydroxylamine head group, *Bioorg. Med. Chem. Lett.* 14 (2004) 449–453.
- [40] M.B. Jeffrey, L. Zuomei, D. Daniel, B. Claire, Methods for specifically inhibiting histone-7 and 8, Patents (2004) US 2004/0072770 A1.
- [41] C. Dizhong, D. Weiping, S. Kand, Y.S. Hong, T.S. Eric, Y. Niefang, Z. Yong, Benzimidazole derivatives: Preparation and pharmaceutical applications, Patents (2007) US 2007/0043043 A1.
- [42] S. Walter, W. Haishan, Y. Zheng, Biaryl linked hydroxamates: Preparation and pharmaceutical applications, Patents (2007) US 2007/0167499 A1.
- [43] L. Ze-Yi, W. Haishan, Z. Yan, Aclyurea connected and sulfonamide connected hydroxamates, Patents (2008) US 2008/0070954 A1.
- [44] J.B. Joseph, B. Sriram, Uses of selective inhibitors of HDAC8 for treatment of T-cell proliferative disorders, Patents (2008) US 2008/0112889 A1.
- [45] B.R. Brooks, R.E. Brucoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem.* 4 (1983) 187–217.
- [46] S. Daniela, L. Christian, M.S. Theodora, P. Anja, W.H. Rolf, L. Thierry, Pharmacophore modeling and in silico screening for new P450 19 (aromatase) inhibitors, *J. Chem. Inf. Model.* 46 (2006) 1301–1311.
- [47] N. Bharatham, K. Bharatham, K.W. Lee, Pharmacophore identification and virtual screening for methionyl-tRNA synthetase inhibitors, *J. Mol. Graph. Model.* 25 (2007) 813–823.
- [48] M.A.C. Neves, T.C.P. Dinis, G. Colombo, M.L.S. Melo, An efficient steroid pharmacophore-based strategy to identify new aromatase inhibitors, *Eur. J. Med. Chem.* 44 (2009) 4121–4127.
- [49] J. Bostrom, Reproducing the conformations of protein-bound ligands: A critical evaluation of several popular conformational searching tools, *J. Comput. Aided Mol. Des.* 15 (2001) 1137–1152.
- [50] M. Vieth, J.D. Hirst, C.L. Brooks, Do active site conformations of small ligands correspond to low free-energy solution structures? *J. Comput. Aided Mol. Des.* 12 (1998) 563–572.
- [51] J.A. Yankeelov Jr., D.E. Koshland Jr., Evidence for conformation changes induced by substrates of phosphoglucomutase, *J. Biol. Chem.* 240 (1965) 1593–1602.
- [52] J. Kirchmair, C. Laggner, G. Wolber, T. Langer, Comparative analysis of protein-bound ligand conformations with respect to catalyst's conformational space subsampling algorithms, *J. Chem. Inf. Model.* 45 (2005) 422–430.
- [53] M.L. Verdonk, J.C. Cole, M.J. Hartshorn, C.W. Murray, R.D. Taylor, Improved protein-ligand docking using GOLD, *Proteins* 52 (2003) 609–623.
- [54] T. Langer, G. Wolber, Pharmacophore definition and 3D searches, *Drug Discov. Today Technol.* 1 (2004) 203–207.
- [55] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Delivery Rev.* 46 (2001) 3–26.
- [56] F. Cheng, Q. Wang, M. Chen, F.A. Quiocho, J. Ma, Molecular docking study of the interactions between the thioesterase domain of human fatty acid synthase and its ligands, *Proteins* 70 (2008) 1228–1234.
- [57] R. Thomsen, M.H. Christensen, MolDock: A new technique for high-accuracy molecular docking, *J. Med. Chem.* 49 (2006) 3315–3321.
- [58] K. Boda, T. Seidel, J. Gasteiger, Structure and reaction based evaluation of synthetic accessibility, *J. Comput. Aided Mol. Des.* 21 (2007) 311–325.
- [59] A. Zaliani, K. Boda, T. Seidel, A. Herwig, C.H. Schwab, J. Gasteiger, H. Claußen, C. Lemmen, J. Degen, J. Pärn, M. Rarey, Second-generation de novo design: A view from a medicinal chemist perspective, *J. Comput. Aided Mol. Des.* 23 (2009) 593–602.
- [60] A.B. Wagner, SciFinder Scholar 2006: An empirical analysis of research topic query processing, *J. Chem. Inf. Model.* 46 (2006) 767–774.
- [61] Y. Wang, E. Bolton, S. Dracheva, K. Karapetyan, B.A. Shoemaker, T.O. Suzek, J. Wang, J. Xiao, J. Zhang, S.H. Bryant, An overview of the PubChem BioAssay resource, *Nucleic Acids Res.* 38 (2010) D255–266.
- [62] J.R. Somoza, R.J. Skene, B.A. Katz, C. Mol, J.D. Ho, A.J. Jennings, C. Luong, A. Arvai, J.J. Buggy, E. Chi, J. Tang, B.C. Sang, E. Verner, R. Wynands, E.M. Leahy, D.R. Dougan, G. Snell, M. Navre, M.W. Knuth, R.V. Swanson, D.E. McRee, L.W. Tari, Structural snapshots of human HDAC8 provide insights into the Class I histone deacetylases, *Structure* 12 (2004) 1325–1334.
- [63] D.P. Dowling, S.L. Gantt, S.G. Gattis, C.A. Fierke, D.W. Christianson, Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors, *Biochemistry* 47 (2008) 13554–13563.